Tatouage asymétrique

Teddy Furon

Tatouage de documents audiovisuels numériques

Traité IC2 - HERMES

version 0.1

3 juillet 2002

Table des matières

Notations

Liste de notations pour le traité IC2 sur le tatouage. Cette liste a pour but d'homogénéiser les notations ; certaines notations propres à certaines parties ne sont pas indiquées, et laissées au choix des rédacteurs.

Pour toutes remarques, merci de contacter Franck Davoine et Stéphane Pateux.

0.1. Tatouage

$\mathbf{m}$	Message (m_j : j^{eme} bit du message $\mathbf{m}$)
$\mathbf{w}$	Marque
$\mathbf{u}$	Mot de code
$\mathbf{s}$	Signal original (s_i : i^{eme} bit du message $\mathbf{m}$) *(on prendra l'indice i pour le signal, et l'indice j pour le message).*
$\mathbf{r_1}$	Signal marqué
$\mathbf{r_2}$	Signal attaqué
$\mathbf{b}$	Bruit
α	Force de marquage
$\sigma^2_{\mathbf{S}}$	Variance du signal $\mathbf{s}$
D_1	Distorsion causée par l'insertion

D_2	Distorsion causée par l'attaque
Θ_k	Clé
L	Taille du signal hôte
K	Taille du message
N	Taille du message codé
R	Rendement du codeur ($R = \frac{K}{N}$)
Q	Fonction Q : $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty exp(\frac{-u^2}{2})du$
$erfc$	Fonction d'erreur complémentaire : $erfc = \frac{2}{\sqrt{\pi}} \int_x^\infty exp(-u^2)du = 2Q(x\sqrt{2})$
C	Capacité
H	Entropie
I	Information mutuelle
p	Probabilité
p_{fa}	Probabilité de fausse alarme
p_e	Probabilité d'erreur
p_b	Probabilité d'erreur bit
η	Seuil
$\mathcal{N}(\mu, \sigma^2)$	Distribution Gaussienne (normale)
d_{GG}	Distribution Gaussienne Généralisée : $d_{GG}(x) = \frac{\nu\beta^{\frac{1}{\nu}}}{2\Gamma(\frac{1}{\nu})}\ exp(-\beta\|x - x_0\|^\nu)$.
d_W	Distribution de Weibull : $d_W(x) = \frac{c}{\alpha}x^{c-1}\ exp(-\frac{x^c}{\beta})$, avec $x, c, \alpha > 0$
$\mathcal{H}$	Hypothèses : le signal est tatoué ($\mathcal{H}_1$), le signal n'est pas tatoué ($\mathcal{H}_0$)

0.2. Opérateur - Mesures - Unités

$\otimes$	Multiplication terme à terme : $[X \otimes Y]_i = x_i \times y_i$
$*$	Convolution
$\cdot$	Produit scalaire : $X \cdot Y = \sum_i X_i Y_i$
c	Corrélation : $c(X, Y) = X \cdot Y$
$exp(.)$	Exponentiel
EQM	Erreur Quadratique Moyenne
EQMP	Erreur Quadratique Moyenne Pondérée
PSNR	Peak Signal over Noise Ratio : mesure de qualité d'un signal codé sur 8 bits ; cette mesure est définie par $PSNR = -10 \log_{10} \frac{EQM}{255^2}$
SNR	Signal over Noise Ratio : rapport signal à bruit.
dB	Décibel. Unité du PSNR.
bpp	Bits par points (bits par échantillon - signal ou image). Coût de codage ramené à un échantillon.
$kbits/s$	Kilo-bits par secondes. Unité de débit. (également $bits/s$ pour bits par secondes).
t	temps.

0.3. Divers

$\mathbf{X}$	Variable aléatoire.
$\mathbf{x}$	Occurrence de la v.a. $\mathbf{X}$.
$\mathbb{E}$	Espérance mathématique. *(On rappelera bien sur quoi est calculée l'espérance, sur quelle hypothèse)*.
$\|A\|$	Cardinal de A si A est un ensemble ou valeur absolue si A est un nombre.
Θ	Paramètres généraux.
$V(\Theta)$	Vraissemblance de Θ.
h	Réponse impulsionnelle du filtre h (indices notés h_i).

$\mathcal{I}$	Image
$\mathcal{S}$	Signal

Chapitre 1

Tatouage asymétrique

1.1. Introduction

Dans ce chapitre, le concept de tatouage asymétrique est introduit de façon simple et pédagogique. Tout d'abord, deux applications se servant du tatouage numérique dans un système de protection sont brièvement décrites. On suppose qu'il existe une technique de marquage invisible et robuste satisfaisant le cahier des charges de chacune de ces applications.

Cependant, des doutes naissent quant à la garantie qu'un pirate ne pourra pas 'hacker' le système de protection. Effectivement, une analyse des menaces révèlent que l'invisibilité et la robustesse du marquage ne sont pas des mesures suffisantes pour garantir la pérennité du système. Un nouveau critère est défini : la sécurité. Cette notion a déjà été introduite dans la section 6.5.3.

Le tatouage asymétrique est alors présenté comme une méthode donnant des niveaux de sécurité plus élevés et convenant mieux aux études de cas du début de chapitre. L'idée de base du tatouage asymétrique et la méthode la plus connue pour la mettre en œuvre sont détaillées. Le chapitre se conclut sur une comparaison des performances avec l'étalement de spectre à séquence directe.

1.2. Etudes de cas

Cette section présente deux applications utilisant le tatouage numérique. Elle pose des questions quant à la fiabilité de cette technologie dans leur contexte applicatif.

Chapitre rédigé par Teddy FURON.

1.2.1. *Protection de copie*

Un système de protection de copie est un circuit embarqué dans les appareils d'électronique grand public de lecture ou écriture d'un certain support. Il contrôle la légitimité de laisser l'utilisateur enregistrer un contenu multimédia. Par exemple, des enregistreurs de DVD de salon contiennent des systèmes de protection de copie. Les architectures complexes de ces systèmes utilisant des primitives cryptographiques et le tatouage numérique, ne sont pas détaillées dans ce livre. Pour faire court, le tatouage numérique y est, en général, utilisé pour distinguer les contenus libres de droit comme les vidéos personnelles de l'utilisateur des contenus protégés car copyrightés. Ainsi, ces derniers sont marqués dans les studios de production avant d'être stockés sur un support et commercialisés. Le détecteur de tatouage embarqués dans un enregistreur de salon empêche la copie dès qu'il trouve la présence de la marque dans un contenu.

Dans cette application, il s'agit de cacher la présence de la marque et non de communiquer un message secret. L'appareil n'extrait pas un message caché mais détecte la présence d'une marque (cf. section 1.6.4.). Il est clair que le système de protection de copie répond à un standard particulier. Celui-ci indique la seule et unique technique de tatouage à utiliser pour être conforme à la norme. Il existe une seule clé secrète, utilisée pour marquer tous les contenus à protéger et enfouie dans tous les appareils compatibles avec ce standard.

Ce manque de diversité dans les clés et dans les messages à cacher n'aide-t-il pas les pirates à casser le système de protection de copie ? En effet, celui-ci a quasiment à sa disposition une infinité de contenus marqués par le même algorithme et avec la même clé. Même si la marque est un signal très faible noyé dans les signaux originaux, est-on sûr que le pirate ne puisse pas extraire de cette infinité de signaux marqués de l'information concernant celle-ci. De plus, est-il prudent de laisser le détecteur de tatouage, même enfoui dans une puce de silicium, à la portée des pirates ? Est-on sûr que le pirate ne puisse pas extraire de l'information concernant la clé secrète en observant le comportement de l'enregistreur de salon (refus ou acceptation de copier) pour différentes entrées.

1.2.2. *Web Spider*

Un web-spider est un ordinateur naviguant sur Internet à la recherche de contenus spécifiques. Il s'agit, par exemple, d'images créées par un même artiste. Ces images sont défendues par le droit d'auteur. Le web-spider recherche des sites utilisant ces images. Il prévient alors l'artiste qui vérifie que ces derniers ont bien demandé son autorisation.

Pour ce faire, on peut imaginer que l'auteur ait tatoué ses images avec sa clé secrète. Ainsi, le web-spider n'a plus qu'à télécharger les images des sites et y détecter

l'éventuelle présence de la marque de l'auteur. Contrairement à la première application, le détecteur de tatouage n'est plus à la portée des pirates. Cependant, toutes les images d'un même auteur sont tatouées par une même clé secrète. Un pirate peut-il tirer profit de cette information ?

1.2.3. *Conclusion*

Ces deux applications ont des cahiers des charges similaires.

– Aucune dégradation perceptive pour respecter la haute qualité des contenus,

– Simple détection d'une présence ou absence de marque,

– Robustesse à des attaques classiques,

– Faible complexité de l'algorithme de détection,

– Sécurité relative.

Le cinquième critère est pour l'instant assez subjectif. Par exemple, le comité de normalisation du système de protection de copie des DVD vidéo souhaite une technique qui ne soit pas trivialement attaquable[1]. En revanche, le comité de normalisation du système de protection de copie pour la musique numérique (i.e. le SDMI) recommandait une technique de tatouage éprouvée. C'est pourquoi, avant de terminer son processus de normalisation, il a lancé un challenge où les experts étaient invités à chercher des failles de sécurité. En quelques semaines, deux équipes de recherche ont 'cassé' les quatre techniques de tatouage pressenties. Cela montre que le concept de sécurité est encore méconnue des concepteurs de tatouage numérique [KAL 01]

Les deux études de cas soumettent l'idée qu'une technique robuste n'assure pas obligatoirement qu'un système de protection est infaillible et inattaquable. Ainsi un critère de sécurité est nécessaire. La sécurité mesure l'impact d'un traitement purement intentionnel dédié à enlever la marque. On parle aussi d'attaque malicieuse au sens où l'attaquant connaît parfaitement l'algorithme d'incrustation du tatouage dans lequel il cherche une faille.

1.3. Analyse de la sécurité

Dans cette section, la cryptographie, science bien plus mature que le tatouage numérique, aide à définir le concept de sécurité et à donner une mesure de ce critère.

1. 'Keep Honest People Honest' est la devise de ce comité.

1.3.1. *Principe de Kerckhoffs*

Le principe fondateur de la cryptographie a été établi en 1883 par A. Kerckhoffs [KER 83]. Il stipule que l'inventeur d'une technique de chiffrement doit supposer que l'attaquant connaît tout de l'algorithme excepté un paramètre secret. Ainsi, la sécurité du crypto-système doit reposer uniquement sur la mise au secret de cette clé, l'algorithme étant public.

Le corollaire de ce principe est qu'une attaque malicieuse se décompose en deux parties. Initialement, le pirate connaît les algorithmes du système de protection. Il ne lui manque que les paramètres secrets pour pirater aisément et à moindre frais. En effet, avec cette connaissance, il est relativement facile de trouver le changement le moins perceptible qui 'lessive' la marque. Dans un premier temps, l'adversaire observe des données publiques. Il essaie d'estimer les clés secrètes à partir de ces observations. La deuxième étape consiste à porter l'attaque sur des contenus en tirant profit du savoir des clés secrètes estimées précédemment.

La différence entre robustesse et sécurité est maintenant plus marquée. La robustesse mesure l'impact sur l'extraction du marquage de transformations appliquées de façon intentionnelle ou non sur le contenu. Ces transformations sont des traitements usuelles comme ceux évoqués dans la section 3.4.2.La sécurité n'est concernée que par les attaques où l'adversaire tire profit des fuites d'information concernant la clé secrète pour mener son attaque.

Pour donner un cadre mathématique à ce concept, nommons l'algorithme A et la clé secrète Θ_k. Au début du jeu, l'adversaire ignore tout de la clé secrète qui a été tirée au hasard par le marqueur. Cette ignorance se mesure par l'entropie de la clé $H(\Theta_k)$. Puis, des contenus marqués sont fabriqués et rendus publiques. Soit $\{O^{(1)}, \ldots, O^{(n)}\}$ la suite d'observations faites par l'adversaire. La fuite d'information se mesure d'après Shannon par l'information mutuelle :

$$I(O^{(1)}, \ldots, O^{(n)}; \Theta_k | A) \qquad [1.1]$$

Ainsi, l'ignorance de l'adversaire tend à diminuer au fur et à mesure qu'il accumule les observations.

$$H(\Theta_k | O^{(1)}, \ldots, O^{(n)}) = H(\Theta_k) - I(O^{(1)}, \ldots, O^{(n)}; \Theta_k | A) \qquad [1.2]$$

Lorsque l'entropie conditionnelle $H(\Theta_k | O^{(1)}, \ldots, O^{(n)})$ est proche de zéro, l'adversaire a réuni suffisamment d'observations pour connaître la clé secrète. Le niveau de sécurité se définit ainsi par ce nombre d'observations nécessaires. Un système de protection est dit parfait si aucune information ne 'fuit' des observations

publiques : $I(O^{(1)}, \ldots, O^{(n)}; \Theta_k | A) = 0$. Alors, l'ignorance de l'adversaire reste constante quelque soit le nombre d'observations.

Ces définitions ont été introduites en cryptographie par C.E. Shannon en 1949 [SHA 49]. Leur adaptation au monde du tatouage numérique est récente [FUR 02a]. Cependant, si les variables aléatoires sont discrètes en cryptographie, on aura affaires à des signaux en tatouage, i.e. à des variables aléatoires continues. Les calculs d'entropie sont alors à interpréter avec précaution puisqu'ils ne correspondent plus à des quantités d'information.

1.3.2. *Classes d'attaques potentielles*

Il est clair qu'un élément de première importance est le type d'observations dont dispose le pirate. Cela dépend de l'application pour laquelle le tatouage est utilisé. Sachant l'algorithme de marquage (d'après le principe de Kerckhoffs), l'attaquant observant une image marquée $\mathcal{I}$, a en fait accès au vecteur $\mathbf{r_1}$. Une classification en types d'attaques est établie pour simplifier les études de sécurité. Il existe de nombreuses classes d'attaques, les plus évidentes ont été citées ci-dessous.

– *Attaque à contenu tatoué.* L'adversaire observe uniquement des contenus marqués. Autrement dit, il possède une collection de signaux marqués $O^{(i)} = \mathbf{r_1}^{(i)}$.

– *Attaque à paire de contenus original-tatoué.* L'adversaire a accès à des paires de contenus. Une paire est constituée d'une version originale du contenu (en plus ou moins bonne qualité) et de sa version marquée. Autrement dit, une observation est constituée d'une paire $O^{(i)} = (\mathbf{s}^{(i)}, \mathbf{r_1}^{(i)})$.

– *Attaque à contenu original choisi.* L'adversaire a accès à un marqueur de contenus. C'est une boîte noire scellée renfermant une clé secrète. Le pirate y insère des contenus originaux en entrée, et observe les contenus marqués résultants. Autrement dit, c'est une attaque à paire de contenus original-tatoué où l'adversaire a le choix dans les contenus.

– *Attaque à contenu tatoué choisi* ou attaque de l'oracle. L'adversaire a accès à un détecteur de contenus. C'est une boîte noire scellée renfermant une clé secrète. Le pirate y insère des contenus en entrée et observe en sortie les messages décodés : $O^{(i)} = (\mathbf{r_1}^{(i)}, \hat{\mathbf{m}}^{(i)})$

1.3.3. *Exemple concret*

Dans l'application protection de copie de la section 1.2.1, outre l'attaque à contenu tatoué, l'attaque à paire de contenu original-tatoué est possible. Un pirate pourrait retrouver une version originale d'un film ancien remasterisé, marqué et commercialisé. Ou encore, les bandes annonces sur Internet des films actuellement en salle seront des éléments de vidéo originales des films DVD qui sortiront dans un an. L'attaque à

contenu tatoué choisi est aussi possible puisqu'un enregistreur de DVD contiendra un détecteur de marque. Pour l'application Web-spider, seule l'attaque à contenu tatoué est de mise.

Un deuxième critère important est la portée de l'attaque. Dans la protection de copie, il n'existe qu'un seul système de protection, donc qu'une seule clé secrète. La découverte de cette clé est un 'hack' total de l'application. Ainsi, si un pirate se sert de bande-annonces, ce n'est pas pour pirater ces films mais tous les films protégés par ce système. Dans l'application Web-spider, une clé est associée à un ayant-droit. La découverte d'une clé permet de pirater facilement les contenus de l'ayant-droit associé, mais les autres contenus sont toujours protégés.

La technique de tatouage choisie pour le système de protection de copie des futurs DVD est une technique à étalement de spectre sur les pixels des images. La marque a une structure propre si bien que la technique de tatouage est robuste à une translation et un changement d'échelle de l'image (cf. section 2.5.2.3.).On suppose une technique de marquage simpliste : $\mathbf{r_1} = \mathbf{s} + \mathbf{w}$. La clé secrète est la marque $\mathbf{w}$. La détection se fait par une corrélation (cf. Eq. [1.36]).

Supposons que le pirate ait accès aux signaux originaux de plusieurs vidéos très différentes. On suppose alors que les signaux sont statistiquement indépendants. En moyennant n signaux marqués, le signal $\bar{\mathbf{r_1}}$ constitue une estimation de la marque :

$$\bar{\mathbf{r_1}} = \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{r_1}^{(i)} = \mathbf{w} + \frac{1}{n}\sum_{i=0}^{n-1} \mathbf{s}^{(i)} = \mathbf{w} + \bar{\mathbf{s}} \qquad [1.3]$$

La puissance du signal $\bar{\mathbf{s}}$ décroît en $1/n$ vers zéro. Pour $n \sim \sigma_{\mathbf{s}}^2/\sigma_{\mathbf{w}}^2$, $\bar{\mathbf{r_1}}$ est une estimation du signal secret avec un bruit résiduel de puissance égale. C'est un rapport signal à bruit convenable pour une attaque. En effet, le pirate ne cherche pas à retrouver l'image originale. Il désire juste créer une image de qualité acceptable et considérée comme non tatouée par le détecteur. Autrement dit, il cherche à rendre la corrélation de l'équation (1.36) inférieure au seuil η. On dira que le niveau de sécurité de l'étalement de spectre contre une attaque à contenu tatoué est de l'ordre de $O(\sigma_{\mathbf{s}}^2/\sigma_{\mathbf{w}}^2)$. En règle général, ce rapport de puissance est de l'ordre de -20 dB. Il faut donc obtenir une centaine de contenus marqués différents pour obtenir une bonne estimation de la clé. Pour l'application de protection de copie des DVD, il est clair que cette attaque constitue une menace réelle. Pour l'application Web-Spider, il sera conseillé de changer régulièrement de clé si la collection d'images d'un auteur est de l'ordre de la centaine. Ceci rend le détecteur plus complexe puisqu'il doit vérifier la présence de plusieurs marques. Cela augmente aussi le risque de fausses alarmes. D'autres évaluation de niveau de sécurité pour des techniques différentes sont données dans [FUR 02b].

1.4. Tatouage asymétrique

Jusqu'en 1999, toutes les techniques de tatouage étaient symétriques (aussi appelées 'à clé privée'). La symétrie signifie que le processus de détection utilise les mêmes paramètres secrets que le processus d'incrustation. Ceci a été illustré à la section précédente où le signal de tatouage $\mathbf{w}$ était sensé jouer le rôle de clé secrète connue à la fois du marqueur et du détecteur. Le concept d'asymétrie a été inventé pour palier les attaques survenant lorsque la diversité des clés secrètes et des messages à cacher est extrêmement faible comme c'est la cas pour les deux études de cas de la section 1.2.

1.4.1. *Principe de base*

Une diversité des clés secrètes et des messages à cacher faible voire nulle se traduit mathématiquement par des entropies $H(\Theta_k)$ et $H(\mathbf{m})$ presque nulle. Le signal de tatouage est alors une fonction déterministe des contenus d'entrée. L'idée de base est ici de recréer artificiellement de l'entropie et de rendre le marquage des contenus aléatoire. Pour cela, la marque $\mathbf{w}$ est non seulement fonction de la clé secrète mais aussi d'une variable aléatoire a appelé *alea*. Ainsi deux versions marquées d'un même contenu original seront différentes car l'*alea* change à chaque marquage. La marque $\mathbf{w}$ est ainsi jamais la même.

La difficulté est maintenant de construire le détecteur. En effet, lorsque le contenu reçu est tatoué, le détecteur ne connaît pas la valeur de l'*alea* prise lors de son marquage. D'où le terme 'asymétrie' : les paramètres servant au marquage ne sont pas identiques à ceux du détecteur. Il faut donc trouver une propriété statistique capable de repérer la présence de la marque sans pour autant la connaître. Le test de détection est alors un test non-paramétrique. La méthode présentée ci-dessous est basé sur des statistiques d'ordre deux (i.e. fonction d'autocorrélation, densité spectrale de puissance).

1.4.2. *Références*

Plusieurs méthodes asymétriques, inventées de façon indépendante, ont été proposées. R. Van Schyndel et A. Tirkel ont présenté leur idée dans l'article [SCH 99] qui a été analysée [EGG 99] et améliorée [J.E 00] par J. Eggers et B. Girod. La proposition de J. Smith et C. Dodge est publiée dans les actes du troisième workshop 'Information Hiding' [SMI 99]. Elle a été redécouverte par G. Sylvestre et N. Hursley [SIL 01] d'une part, et J. Stern et J.-P. Tillich [STE 01] avec une approche plus cryptographique, d'autre part. Enfin, l'auteur et P. Duhamel ont exposé leur méthode au même workshop [FUR 99]. Nous avons récemment prouvé qu'en fait toutes ces méthodes, en apparence très différentes, sont basées sur une formulation mathématique de l'algorithme de détection commune [FUR 01]. Nous présentons dans la suite de l'article la méthode exposée dans l'article [FUR 03].

1.4.3. *Algorithmes*

Dans cette méthode asymétrique, le détecteur ne compare pas le vecteur extrait $\mathbf{r_1}$ à un signal de tatouage spécifique $\mathbf{w}$, mais il vérifie si $\mathbf{r_1}$ a une propriété statistique due à la présence de $\mathbf{w}$. Nous décrivons d'abord l'algorithme d'incrustation.

Tout d'abord, un signal $\mathbf{v}$ dont les composantes sont i.i.d. représentant un processus aléatoire centré blanc Gaussien de puissance $\sigma^2_{\mathbf{v}}$ est créé grâce à l'*alea* a. Puis il est convolué par le filtre h. Le signal résultant est alors entrelacé. Cette entrelaceur agit comme une permutation pseudo-aléatoire π de $\{0,\ldots,L-1\}$ dans $\{0,\ldots,L-1\}$. Elle brasse ainsi les échantillons du signal. Les signaux entrelacés seront accentués avec le symbole tilde : $\widetilde{s}_i = s_{\pi(i)}$ et $r_{1i} = \widetilde{r}_{1\pi^{-1}(i)} \quad \forall i \in \{0,\ldots,L-1\}$. Finalement, la marque est définie par : $w_i = (h * \mathbf{v})_{\pi(i)} \quad \forall i \in \{0,\ldots,L-1\}$. La formule d'incrustation est donc :

$$r_{1i} = s_i + \alpha(h * \mathbf{v})_{\pi(i)} \quad \forall i \in \{0,\ldots,L-1\}$$

Le filtre normalisé h et la permutation π sont les paramètres secrets du processus d'incrustation. Ils ne changent pas la puissance des signaux, donc $\sigma^2_{\mathbf{w}} = \sigma^2_{\mathbf{v}}$. Noter que n'importe quel signal $\mathbf{v}$ pseudo-aléatoire centré blanc Gaussien convient. Grâce à l'*alea*, il est tiré de façon aléatoire à chaque marquage. La détection n'a pas besoin de la connaissance du signal $\mathbf{v}$. Elle doit connaître le désentrelaceur (c'est à dire la permutation inverse π^{-1}) et le module de la réponse fréquentielle du filtre h. Cet ensemble de paramètres $\{\pi^{-1}, |H(f)|\}$ caractérise la propriété statistique attendue : le spectre du signal entrelacé $\widetilde{\mathbf{r_1}}$ à la forme de $|H(f)|^2$. Un simple test décide à quelle hypothèse ($\mathcal{H}_0$ ou $\mathcal{H}_1$) le contenu reçu appartient le plus vraisemblablement.

• $\mathcal{H}_0$: Le signal reçu $\mathbf{r}$ n'est pas tatoué, donc il ne partage pas la propriété statistique particulière. Grâce à l'action supposée idéale de la permutation pseudo-aléatoire π^{-1}, les échantillons de $\widetilde{\mathbf{r}}$ sont supposés représenter un processus blanc et stationnaire, si bien que son spectre $S_0(f)$ est constant : $S_0(f) = \sigma^2_{\mathbf{r}} + \mu^2_{\mathbf{r}}\delta(f)$ où $\mu_{\mathbf{r}}$ et $\sigma^2_{\mathbf{r}}$ sont la moyenne et la variance des échantillons de $\mathbf{r}$.

• $\mathcal{H}_1$: Le signal extrait $\mathbf{r}$ a été tatoué. Comme $\widetilde{\mathbf{w}}$ et $\widetilde{\mathbf{r_1}}$ sont des signaux aléatoires statistiquement indépendants et stationnaires, la relation suivante donne :

$$\begin{aligned} \varphi_{\widetilde{\mathbf{r}}}[l] &= \mathbb{E}_{\mathrm{R}}\{\tilde{\mathrm{r}}[\mathrm{i}].\tilde{\mathrm{r}}[\mathrm{i}+\mathrm{l}]\} = \varphi_{\widetilde{\mathbf{s}}}[\mathrm{l}] + \varphi_{\widetilde{\mathbf{w}}}[\mathrm{l}] \\ \Phi_{\widetilde{\mathbf{r}}}(f) &= S_1(f) = \Phi_{\widetilde{\mathbf{s}}}(f) + \Phi_{\widetilde{\mathbf{w}}}(f) \end{aligned}$$

où $\varphi_{\widetilde{\mathbf{r}}}[.]$ la fonction de corrélation du signal $\widetilde{\mathbf{r}}$ et $\Phi_{\widetilde{\mathbf{r}}}(.)$ sa transformée de Fourrier, qui est la densité spectrale de puissance (hypothèse de stationnarité). Comme $\widetilde{\mathbf{w}} = h * \mathbf{v}$,

$\Phi_{\widetilde{\mathbf{w}}}(f) = |H(f)|^2$. Le filtre h est normalisé si bien que $\int |H(f)|^2 \, df = 1$. Finalement, la densité spectrale de puissance espérée dans le cas $\mathcal{H}_1$, est la suivante :

$$\begin{aligned} S_1(f) &= \sigma_{\mathbf{s}}^2 + \sigma_{\mathbf{w}}^2 + \mu_{\mathbf{s}}^2 \delta(f) + \sigma_{\mathbf{w}}^2 |H(f)|^2 & [1.4] \\ &= \mu_{\mathbf{r}}^2 \delta(f) + \sigma_{\mathbf{r}}^2 + \sigma_{\mathbf{w}}^2 (|H(f)|^2 - 1) & [1.5] \end{aligned}$$

Ainsi, le spectre $S_1(f)$ est de la forme de $|H(f)|^2$.

La détection doit distinguer les signaux au spectre plat comme $S_0(f)$ de ceux dont le spectre est profilé comme $S_1(f)$. Le test effectué sur les signaux entrelacés par π^{-1} ne dépend donc que du gabarit $|H(f)|^2$. On choisit un test d'hypothèses en analyse spectrale basée sur un critère de maximum de vraisemblance :

$$\hat{m} = \begin{cases} 1 & \text{si} \quad v = V_L(\widetilde{\mathbf{r}}, S_0) - V_L(\widetilde{\mathbf{r}}, S_1) \geq \eta \\ 0 & \text{sinon} \end{cases} \qquad [1.6]$$

où η est un seuil positif dépendant de la probabilité de fausse alarme p_{fa} fixée dans le cahier des charges et $V_L(\widetilde{\mathbf{r}}, S_i)$ est la partie principale de la vraisemblance de Whittle que le spectre du processus aléatoire $\widetilde{\mathbf{r}}$ corresponde avec la densité spectrale de puissance S_i. Son expression simplifiée est la suivante :

$$V_L(\widetilde{\mathbf{r}}, S_i) = 2L \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{I_L(f)}{S_i(f)} + \log S_i(f) df \qquad [1.7]$$

où $I_L(f)$ est le périodogramme du signal $\widetilde{\mathbf{r}}$:

$$I_L(f) = \left| \sum_{k=0}^{L-1} \widetilde{r}[k].e^{2\pi i k f} \right|^2 \quad \forall f \in]-\frac{1}{2}, \frac{1}{2}]$$

1.4.4. *Performances*

1.4.4.1. *Versatilité*

Le signal $\mathbf{w}$ enfoui dans le contenu est un bruit Gaussien blanc grâce à l'action de l'entrelaceur π. Ainsi, cette méthode s'adapte à toute technique de tatouage à étalement de spectre. Seule la manière de créer le signal de tatouage et de détecter sa présence ont changé. Au lieu de choisir sans cesse le même vecteur, on tire, pour chaque contenu, un signal pseudo-aléatoire que l'on filtre puis entrelace. Le lecteur trouvera dans [FUR 03] l'adaptation de l'algorithme de détection lorsque l'équation de marquage est moins simpliste que dans l'exemple traité ci-dessus.

1.4.4.2. *Puissance de détection*

Le cahier des charges des systèmes de protection impose en général un niveau ω de fausse alarme maximum. Le seuil η est calculé de sorte que $p_{fa} < \omega$. On reconnaît ici un test d'hypothèse suivant la stratégie de Neyman-Pearson. Le test est alors d'autant plus efficace que sa puissance $p_p = \mathbb{E}\{\hat{\mathrm{m}}|\mathcal{H}_1\}$ est grande. p_p est en fait une fonction croissante du coefficient de déflection e :

$$e = \frac{\mathbb{E}\{\mathrm{v}|\mathcal{H}_1\} - \mathbb{E}\{\mathrm{v}|\mathcal{H}_0\}}{\sigma_{v|\mathcal{H}_1}} \qquad [1.8]$$

Dans le cas d'une technique de tatouage symétrique à étalement de spectre à séquence directe avec une détection par corrélation, un calcul classique montre que le coefficient de déflection est de l'ordre de :

$$e_s \propto \frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{s}}}\sqrt{L}$$

En revanche, les méthodes asymétriques produisent un coefficient de déflection d'un ordre de grandeur plus faible :

$$e_a \propto \frac{\sigma^2_{\mathbf{w}}}{\sigma^2_{\mathbf{s}}}\sqrt{L}$$

Comme $\sigma_{\mathbf{w}}/\sigma_{\mathbf{s}} < 1$, les méthodes asymétriques sont bien moins efficaces que les techniques symétriques. Pour palier cet inconvénient, la seule solution est d'accroître L. Comme le rapport de puissance $\sigma^2_{\mathbf{w}}/\sigma^2_{\mathbf{s}}$ est de l'ordre de -20 dB, il faut détecter la présence du tatouage sur des signaux environ 10 fois plus grands, ce qui amène de nombreuses difficultés sur la complexité, la taille mémoire et le temps de réponse nécessaires aux détecteurs asymétriques.

1.4.4.3. *Sécurité*

Le rôle de l'entrelaceur est très important. Certes, il blanchit la part du signal provenant du contenu original, ce qui est fondamental pour le bon fonctionnement de la détection. Mais, il joue aussi un rôle essentiel quant à la sécurité de la méthode. En fait, il cache au pirate ce qu'est exactement la propriété statistique attendue des contenus marqués. Sans sa connaissance, le pirate n'a pas accès au domaine où le cœur de la détection a lieu, c'est-à-dire les calculs de vraisemblance de l'Eq. (1.7)). Ainsi, il lui est impossible prédire l'impact de son attaque.

Evaluer le niveau de sécurité de cette méthode asymétrique revient à calculer la complexité nécessaire pour estimer la permutation π. Supposons que le pirate dispose d'une paire de contenus original/tatoué. Il estime le signal de tatouage $\mathbf{w}$ par la différence $\mathbf{r_1} - \mathbf{s}$. L'adversaire doit maintenant trouver la permutation à partir de ce signal

différence. Une possibilité est d'essayer toutes les permutations possibles et de s'arrêter lorsque le signal permuté est coloré. Le deuxième paramètre secret $|H(f)|^2$ et par conséquent, le spectre espéré S_1 ne sont pas connus. Le pirate prendra à la place un test de sphéricité comme celui de Drouiche et Fay [FAY 00], qui détermine la vraisemblance qu'un signal soit coloré quelque soit sa couleur. Cependant, il existe $L!$ permutations possibles. Donnons un ordre de grandeur : Pour $L = 2048$, en utilisant l'approximation de Stirling, $L! \sim 2^{19000}$, ce qui est considérablement plus grand que 2^{280}, nombre de particules de particules dans l'univers !

Cette évaluation du niveau de sécurité est très naïve. Le pirate sait bien que la permutation π^{-1} a de très bonnes propriétés de décorrélation. Or, dans les $L!$ permutations, il existe de nombreuses permutations échangeant uniquement quelques échantillons de place. De plus, il ne souhaite pas trouver π^{-1} exactement ; une permutation suffisamment proche convient. Pour une évaluation correcte de la complexité, nous sommes obligés, à l'instar de A. Kerckhoffs, de dévoiler l'algorithme qui a donné naissance à la permutation π. C'est un générateur de permutations pseudo-aléatoires dont l'entrée est un mot de L_k bits ($L_k << L$), comme on en trouve dans le livre de D. Knuth [KNU 81]. Quelque soit le mot d'entrée, il donne une permutation ayant une bonne action de blanchiment. Parcourir l'espace des permutations revient maintenant à essayer les 2^{L_k} éléments possibles, ce qui est infiniment moins que $L!$ Supposons qu'une génération de permutation suivie d'un test de sphéricité dure une 1μs. L'espérance du temps nécessaire pour tomber sur la bonne permutation est d'environ 2^{L_k-46} années. Cette analyse ne prend pas en compte la loi de Moore, ni la possibilité de paralléliser l'attaque. Elle montre cependant que le niveau de sécurité est nettement supérieur à celui de la méthode symétrique.

Nous pouvons montrer que l'attaque par oracle est de même plus difficile à réaliser avec la méthode asymétrique. La différence majeure est que le détecteur n'est pas linéaire contrairement à ce qui est supposé dans l'attaque détaillée dans l'article [KAL 98]. Sa complexité est en $O(L^2)$ et non en $O(L)$. Certes, ce n'est pas un niveau de sécurité recommandable en cryptographie (i.e. il n'est pas exponentiel), mais notons que le cahier des charges des techniques de tatouage en protection de copie stipule que le détecteur donne une prise de décision toutes les 10 secondes. Or le pirate ne peut pas accélérer ce débit de décision puisqu'il utilise le détecteur comme une boîte noire. Ainsi pour $L = 2048$, $O(L)$ essais prennent environs 6 heures alors que $O(L^2)$ essais prennent un an et demi.

Un autre point critique est d'analyser les menaces pour le système si le pirate découvre, par 'reverse engineering' de l'implémentation, les paramètres $\{\pi^{-1}, |H(f)|\}$. La réponse n'est pas évidente. Le pirate a maintenant accès au signal marqué $\mathbf{r_1}$ et sa version permutée $\widetilde{\mathbf{r_1}}$. Il peut créer un signal permuté $\widetilde{\mathbf{r_2}}$ que le détecteur considérera comme non tatoué. Mais il n'est pas capable de prédire l'impact visuel de cette attaque. De la même manière, il est capable d'utiliser un modèle perceptif dans l'espace

de ‘tatouage’ pour créer un contenu de bonne qualité, mais il ne peut pas prédire l’impact de ce modèle sur la sortie du détecteur. Ceci est dû à l’entrelaceur qui empêche d’avoir accès dans un même domaine au modèle perceptif et à la formule de détection. Cependant, au bout de quelques itérations, l’adversaire parvient en pratique à ses fins. Ainsi, une technique de tatouage asymétrique ne signifie aucunement une technique de détection à clé publique. La clé utilisée côté détecteur doit rester secrète.

1.5. Conclusion

Le tableau 1.1 issu de [FUR 02b] sert de conclusion à ce chapitre. Il compare quelques caractéristiques de l’étalement de spectre par séquence directe et celles d’une méthode asymétrique. Cette dernière offre des niveaux de sécurité supérieurs mais un coefficient de déflection inférieur d’un ordre de grandeur. Le prix à payer pour plus de sécurité est un accroissement de la longueur des signaux à analyser.

Les attaques étudiées dans ce chapitre sont surtout caractéristiques de la protection de copie et du Web-spider. Le manque de diversité des clés secrètes et des messages à cacher engendrent des failles que le tatouage asymétrique se propose de résoudre. Cependant, pour d’autres applications, il n’est pas évident que l’asymétrie soit utile.

D’autre part, l’asymétrie tend à perdre de son intérêt à cause de l’avancée des techniques utilisant l’information de bord. En effet, l’asymétrie consiste avant tout à rendre le marquage plus aléatoire. C’est en partie parce qu’il est plus ’imprévisible’ que le tatouage asymétrique procure des niveaux de sécurité plus forts. Or, dans les techniques tirant profit de l’information de bord, le signal de tatouage est fonction du signal original. Ceci est vrai non seulement à cause du modèle perceptif qui vient moduler mollement la force de tatouage (et ceci quelque soit la technique de tatouage), mais surtout la façon de coder le message dépend de l’information de bord. Cette information de bord est ainsi une source d’entropie. Même si le message à cacher est toujours le même (c’est le cas pour la simple détection d’un présence de marque), le signal de tatouage sera différent d’un contenu original à l’autre. Le grand avantage est que les techniques avec information de bord sont nettement plus efficaces (elles peuvent avoir des coefficients de déflection supérieur à e_s). Cependant, l’attaque à contenu original choisi est leur talon d’Achille [FUR 02b].

On retiendra plutôt que la sécurité est un critère important. Elle s’étudie de la manière suivante :

1) Description de l’application visée et du rôle du tatouage dans le système de protection global.

2) Analyse des menaces possibles sur la technique de marquage.

3) Estimation de la complexité requise pour mener à bien une attaque contre le marquage. Cette complexité définit le niveau de sécurité contre cette attaque.

Critères	DSSS	Asymétrique
coefficient de déflection	$\frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{s}}}\sqrt{L}$	$\frac{\sigma^2_{\mathbf{w}}}{\sigma^2_{\mathbf{s}}}\sqrt{L}$
Niveau de sécurité Attaque à contenu tatoué	$O(\frac{\sigma^2_{\mathbf{s}}}{\sigma^2_{\mathbf{w}}})$	not possible
Niveau de sécurité Attaque à paire original-tatoué	$O(1)$	$O(2^{L_k})$
Niveau de sécurité Attaque à contenu original choisi	$O(1)$	$O(2^{L_k})$
Niveau de sécurité Attaque à contenu tatoué choisi	$O(L)$	$O(L^2)$

Tableau 1.1. *Comparaison étalement de spectre à séquence directe (DSSS) et méthode asymétrique*

4) Impact de l'attaque sur le système de protection global

1.6. Bibliographie

[EGG 99] EGGERS J., B. GIROD, « Robustness of public key watermarking schemes », V^3D^2 *Watermarking Workshop*, Erlangen, Germany, octobre 1999.

[FAY 00] FAY G., Théorèmes limite pour les fonctionnelles de périodogramme, PhD thesis, Ecole Nationale Supérieure des Télécommunications, 2000.

[FUR 99] FURON T., DUHAMEL P., « An Asymmetric Public Detection Watermarking Technique », PFITZMANN A., Ed., *Proc. of the third Int. Workshop on Information Hiding*, Dresden, Germany, Springer Verlag, p. 88-100, septembre 1999.

[FUR 00] FURON T., DUHAMEL P., « Robustness of an asymmetric technique », *Proc. of Int. Conf. on Image Processing*, Vancouver, Canada, IEEE, septembre 2000.

[FUR 01] FURON T., I. VENTURINI, P. DUHAMEL, « Unified approach of asymmetric watermarking schemes », P.W. WONG, E. DELP, Eds., *Security and Watermarking of Multimedia Contents III*, San Jose, Cal., USA, SPIE, 2001.

[FUR 02a] FURON T., J. OOSTVEN, J. VAN BRUGGEN, Security Analysis, Deliverable d.5.5, CERTIMARK IST European Project, 2002.

[FUR 02b] FURON T., Application du tatouage numérique à la protection de copie, PhD thesis, Ecole Nationale Supérieure des Télécommunications., 2002.

[FUR 03] FURON T., P. DUHAMEL, « An asymmetric watermarking method », *IEEE Trans. on Signal Processing*, vol. 51, n°4, p. 981-995, avril 2003, special issue on signal processing for data hiding in digital media & secure content delivery, IEEE Trans. on Signal Processing.

[J.E 00] J.EGGERS, J.SU, B.GIROD, « Public key watermarking by eigenvectors of linear transforms », *Proc. of the European Signal Processing Conference*, Tampere, Finland, EUSIPCO, septembre 2000.

[KAL 98] KALKER T., « A security risk for publicly available watermark detectors », *Benelux Information Theory Symposium*, May 1998, Veldhoven, The Netherlands.

[KAL 01] KALKER T., « Considerations on watermarking seccurity », *Proc of the IEEE Multimedia Signal Processing workshop*, Cannes, France, p. 201-206, October 2001.

[KER 83] KERCKHOFFS A., « La cryptographie militaire », *Journal des sciences militaires*, vol. 9, p. 5-38, janvier 1883.

[KNU 81] KNUTH D., *The art of computer programming*, Computer Science and Information Processing, Addison-Wesley, 1981.

[SCH 99] SCHYNDEL R. V., A. TIRKEL, I. SVALBE, « Key independent watermark detection », *Int. Conf. on Multimedia Computing and Systems*, vol. 1, Florence, Italy, juin 1999.

[SHA 49] SHANNON C., « Communication theory of secrecy systems », *Bell system technical journal*, vol. 28, p. 656-715, octobre 1949.

[SIL 01] SILVESTRE G., N. HURLEY, G. HANAU, W. DOWLING, « Informed Audio Watermarking using Digtial Chaotic Signals », *Proc. of Int. Conf. on Acoustics, Speech and Signal Processing*, Salt-Lake City, USA, IEEE, mai 2001.

[SMI 99] SMITH J., DODGE C., « Developments in steganography », PFITZMANN A., Ed., *Proc. of the third Int. Workshop on Information Hiding*, Dresden, Germany, Springer Verlag, p. 77-87, septembre 1999.

[STE 01] STERN J., Contribution à la théorie de la protection de l'information, PhD thesis, Université de Paris XI, Orsay, Laboratoire de Recherche en Informatique, mars 2001.